BREAKTHROUGH-ORIENTED RESEARCH PUBLICATION / DATA-RICH HTML MANUSCRIPT

TAIJITU crosses the large-sample threshold in the 171-galaxy profile-free run

A conservative, data-centred synthesis of the TAIJITU programme from the earlier 50-galaxy validation stages to the new 171-galaxy large-sample profile-free run. The manuscript is framed around a precise threshold question: whether the profile-free branch can move from mechanistic plausibility to population-level competitiveness while preserving scientific caution.

Author:Janos Gabor Melegh

Abstract

The TAIJITU programme previously established a two-sector mechanism, a real coupling-driven transition, a resolved numerical sign pathology, and a 50-galaxy applicability-domain result in which inner baryon dominance index (BDI) emerged as the main empirical separator of where the induced dark correction helps. The new Phase 11D v3 run extends the analysis to a 175-object SPARC-like catalogue, of which 171 galaxies were retained after parsing. The strongest new result is that the profile-free branch no longer behaves as a merely diagnostic or qualitative mapping experiment, but crosses a key threshold into population-level competitiveness. On the 171-galaxy run, the profile-free branch preserves a strong BDI–improvement anti-correlation (Spearman $\rho = -0.715$, $p = 4.44\times10^{-28}$), improves over the baryon-only reference in 108/171 galaxies (63%), and slightly outperforms the gNFW reference in 89/171 galaxies (52%).

The proper scientific reading remains deliberately cautious. The absolute χ^2 scale of the 171-galaxy run remains high (median $\chi^2 = 12.04$ for the profile-free branch and 13.51 for the gNFW branch), indicating that the full catalogue is substantially harder than the earlier 50-galaxy validation cohort and that the current modelling pipeline is not yet a universally accurate galaxy-fit theory. This is therefore a threshold result rather than a final one: the low-BDI applicability domain remains visible at much larger sample size, and the profile-free branch is now population-level competitive rather than merely structurally suggestive.

1. Evidence base and scope

This publication integrates four evidence layers: the earlier consolidated TAIJITU manuscript up to the 50-galaxy stage; the 50-galaxy parametric Phase 10e archive and figures; the first 50-galaxy profile-free pilot (Phase 11D v2); and the new 171-galaxy machine-readable Phase 11D v3 run archive. The scientific question is specific and threshold-oriented: does the profile-free branch merely reproduce qualitative trends, or does it become population-level competitive on a large sample while retaining the previously established BDI regime structure?

171 RETAINED GALAXIES	**135/171** MEASURED VFLAT	**-0.715** BDI–PF SPEARMAN P	**108/171** PF BEATS BARYON-ONLY
89/171 PF BEATS GNFW	**12.04** MEDIAN X^2 PROFILE-FREE	**13.51** MEDIAN X^2 GNFW	**89 / 82**

SPECTRAL / POTENTIAL WINNERS

Interpretive boundary. This manuscript does not claim that TAIJITU has already become a universal galaxy-fit theory. It does claim that the profile-free branch has crossed an important threshold: on the supplied 171-galaxy run it is large-sample competitive while preserving a physically interpretable BDI-dependent regime pattern.

2. Comparison with earlier validated stages

The new run should not be interpreted in isolation. Earlier work had already shown, on the 50-galaxy parametric branch, that BDI was the dominant empirical separator and that the low-BDI regime was the natural applicability domain. The first 50-galaxy profile-free pilot recovered the same qualitative pattern but remained quantitatively weaker than the Phase 10e gNFW ceiling branch. The new 171-galaxy run is therefore best read as the next step in a sequence rather than as a disconnected result.

STAGE	SAMPLE	BRANCH	BDI RHO	BDI P	HELP VS BARYON-ONLY	MEDIAN CHI²	REFERENCE COMPARED AGAINST
Phase 10e gNFW branch	50	Parametric gNFW-mapped TAIJITU	-0.789	9.950e-12	19/50 (38%)	1.611	baryon-only
Phase 11D v2 profile-free	50	Profile-free pilot	-0.711	2.456e-06	15/50 (44%)	2.069	baryon-only and Phase 10e
Phase 11D v3 profile-free	171	Profile-free large-sample run	-0.715	4.438e-28	108/171 (63%)	12.04	baryon-only and gNFW

The key comparison is not that all numerical quantities move in the same direction. They do not. Instead, the stable observation across stages is that the BDI signal survives sample growth, while the profile-free branch moves from a merely suggestive pilot at 50 galaxies to a population-level competitive branch on 171 galaxies.

3. Main results of the 171-galaxy run

METRIC	VALUE
Galaxies retained after parsing	171 of 175
Vflat directly measured	135 of 171
Vflat estimated/fallback	36 of 171
BDI–PF Spearman ρ	-0.715
BDI–PF p-value	4.438e-28
PF beats baryon-only	108/171 (63.2%)
PF beats gNFW	89/171 (52.0%)
gNFW beats baryon-only	116/171 (67.8%)

METRIC	VALUE
Median χ^2 baryon-only	53.76
Median χ^2 gNFW	13.51
Median χ^2 profile-free	12.04

> **Central result.** The profile-free branch retains a strong BDI anti-correlation on the 171-galaxy sample and beats the baryon-only reference in a clear majority of galaxies. It also slightly exceeds the gNFW reference in more than half of the sample. Within the present evidence base, that marks a threshold shift from a suggestive pilot to a population-level competitive branch.

This does *not* mean that the problem is solved. The absolute χ^2 level remains high, so the branch is best viewed as scientifically important and promising rather than complete. But it is no longer merely a diagnostic branch: in the supplied large-sample run it crosses into population-level competitiveness.

---

4. The core physical result: BDI remains the primary regime separator

The most important result is that the BDI signal survives the move from the earlier 50-galaxy validation cohort to the 171-galaxy run. The BDI–profile-free improvement relation remains strongly negative, meaning that the systems most helped by the profile-free branch remain those with lower inner baryonic dominance. This is precisely the qualitative structure established earlier on the 50-galaxy parametric branch.

BDI REGIME	171 PF N	171 PF HELP VS BAR	171 PF BEAT GNFW	171 PF MEDIAN X^2	50 PF HELP VS BAR	50 PF MEDIAN X^2	50 GNFW HELP VS BAR	50 GNFW MEDIAN X^2
<0.80	65	96.9%	58.5%	3.687	100.0%	2.319	100.0%	1.552
0.80-0.93	20	80.0%	55.0%	25.99	100.0%	1.998	100.0%	1.583
0.93-0.99	16	56.2%	62.5%	9.796	28.6%	2.059	40.0%	1.614
>=0.99	70	28.6%	42.9%	129.27	12.5%	2.039	4.3%	1.651

Two conclusions follow. First, the low-BDI domain remains the clearest success region: the 171-galaxy run still shows very high help fractions below BDI ≈ 0.93. Second, the ≥0.99 regime remains the hardest, with large χ^2 values and much lower help fractions. The structure is therefore preserved even though the absolute fit quality deteriorates on the larger and more heterogeneous sample.

---

5. Comparison with other scientific baselines

The relevant scientific comparisons in the supplied material are not against vague external claims, but against explicit modelling baselines: baryon-only fits, the earlier parametric gNFW-mapped TAIJITU branch, and the first 50-galaxy profile-free pilot. Against baryon-only, the new profile-free run is clearly stronger than the earlier 50-galaxy parametric help fraction. Against gNFW, the result is more subtle: the profile-free branch wins only slightly in aggregate, and the absolute χ^2 scale shows that both branches struggle on difficult galaxies in the full catalogue.

- **Against baryon-only:** 108/171 profile-free wins, versus 19/50 wins in the earlier 50-galaxy parametric validation.
- **Against gNFW:** 89/171 profile-free wins, compared to a median χ^2 of 12.04 for profile-free and 13.51 for gNFW on the same run.
- **Against the earlier 50-galaxy profile-free pilot:** the new run moves beyond merely recovering the regime structure and becomes a majority-improving branch against baryon-only.

> **Why the χ^2 comparison must be read carefully.** The 171-galaxy run is not directly comparable to the earlier 50-galaxy validation cohort as though they were the same problem. The 171-galaxy catalogue is much broader and evidently much harder; the median χ^2 values of both the gNFW and profile-free branches are therefore far above the 50-galaxy parametric ceiling.

---

6. Mode competition: spectral versus potential mapping

One of the most interesting new results is that the profile-free branch no longer reduces to one dominant mapping mode. The best mode split is nearly even: spectral_v3 wins on 89 galaxies and potential_v3 on 82. This is scientifically important because the earlier 50-galaxy programme logic still treated spectral diagnostics mainly as secondary mechanistic clues behind BDI. On the full run, spectral and potential mappings are both operationally relevant.

BDI REGIME	SPECTRAL_V3 WINNERS	POTENTIAL_V3 WINNERS
<0.80	37	28
0.80-0.93	8	12
0.93-0.99	10	6
>=0.99	34	36

This does not overturn the BDI result. BDI remains the main empirical separator. What it does suggest is that the mechanism-to-observable mapping may require more than a single profile-free operational channel once the sample is large and heterogeneous enough.

---

7. Strongest winners and hardest failures

The most improved galaxies are overwhelmingly low-BDI systems, many of them gas-rich and low-surface-brightness. The worst failures are concentrated in the high-BDI tail, where both profile-free and parametric branches face much more severe modelling difficulty.

Top profile-free winners versus baryon-only

GALAXY	REGIME	INNER_BDI	CHI2_BAR	CHI2_GNFW	CHI2_PF	PF_IMPROVEMENT	DELTA_PF_GNFW	BEST_MODE
UGC12632	<0.80	0.559	97.42	0.900	0.168	578.78	-0.732	spectral_v3
UGC11820	<0.80	0.628	214.88	5.222	0.428	501.56	-4.794	potential_v3
UGC12732	<0.80	0.574	141.08	1.792	0.560	252.09	-1.232	spectral_v3

GALAXY	REGIME	INNER_BDI	CHI2_BAR	CHI2_GNFW	CHI2_PF	PF_IMPROVEMENT	DELTA_PF_GNFW	BEST_MODE
UGC06399	<0.80	0.726	50.64	1.784	0.205	246.94	-1.579	spectral_v3
UGC00191	<0.80	0.692	143.54	12.88	0.633	226.86	-12.25	potential_v3
F568-V1	<0.80	0.514	36.43	0.307	0.202	180.48	-0.105	spectral_v3
UGC05764	<0.80	0.392	223.35	7.806	1.244	179.59	-6.562	potential_v3
F563-V2	<0.80	0.476	20.49	0.241	0.131	156.49	-0.110	spectral_v3
UGC09992	0.80-0.93	0.815	1.804	0.068	0.013	133.83	-0.055	spectral_v3
UGC10310	<0.80	0.650	27.69	0.213	0.229	120.77	0.016	potential_v3

Worst profile-free cases versus baryon-only

GALAXY	REGIME	INNER_BDI	CHI2_BAR	CHI2_GNFW	CHI2_PF	PF_IMPROVEMENT	DELTA_PF_GNFW	BEST_MODE
UGC05253	>=0.99	1.124	149.59	2.433e+03	3.475e+03	0.043	1.041e+03	potential_v3
UGC08699	>=0.99	1.079	48.08	688.52	978.05	0.049	289.53	potential_v3
NGC3726	>=0.99	1.255	25.73	161.20	284.83	0.090	123.63	spectral_v3
UGC03546	>=0.99	1.371	54.05	266.27	539.24	0.100	272.97	potential_v3
NGC3893	>=0.99	1.238	20.92	81.86	195.00	0.107	113.14	potential_v3
UGC06787	>=0.99	1.089	438.38	3.321e+03	3.889e+03	0.113	568.52	potential_v3
NGC3769	>=0.99	1.066	26.79	149.41	232.61	0.115	83.19	potential_v3
UGC06614	>=0.99	1.133	19.03	117.42	161.33	0.118	43.91	potential_v3
UGC03580	>=0.99	1.425	243.22	1.509e+03	1.593e+03	0.153	83.97	potential_v3
UGC02916	>=0.99	1.126	181.55	813.47	1.157e+03	0.157	343.94	potential_v3
NGC3521	>=0.99	1.280	30.91	81.87	196.41	0.157	114.54	spectral_v3
NGC6946	>=0.99	1.265	64.20	261.83	401.88	0.160	40.05	potential_v3

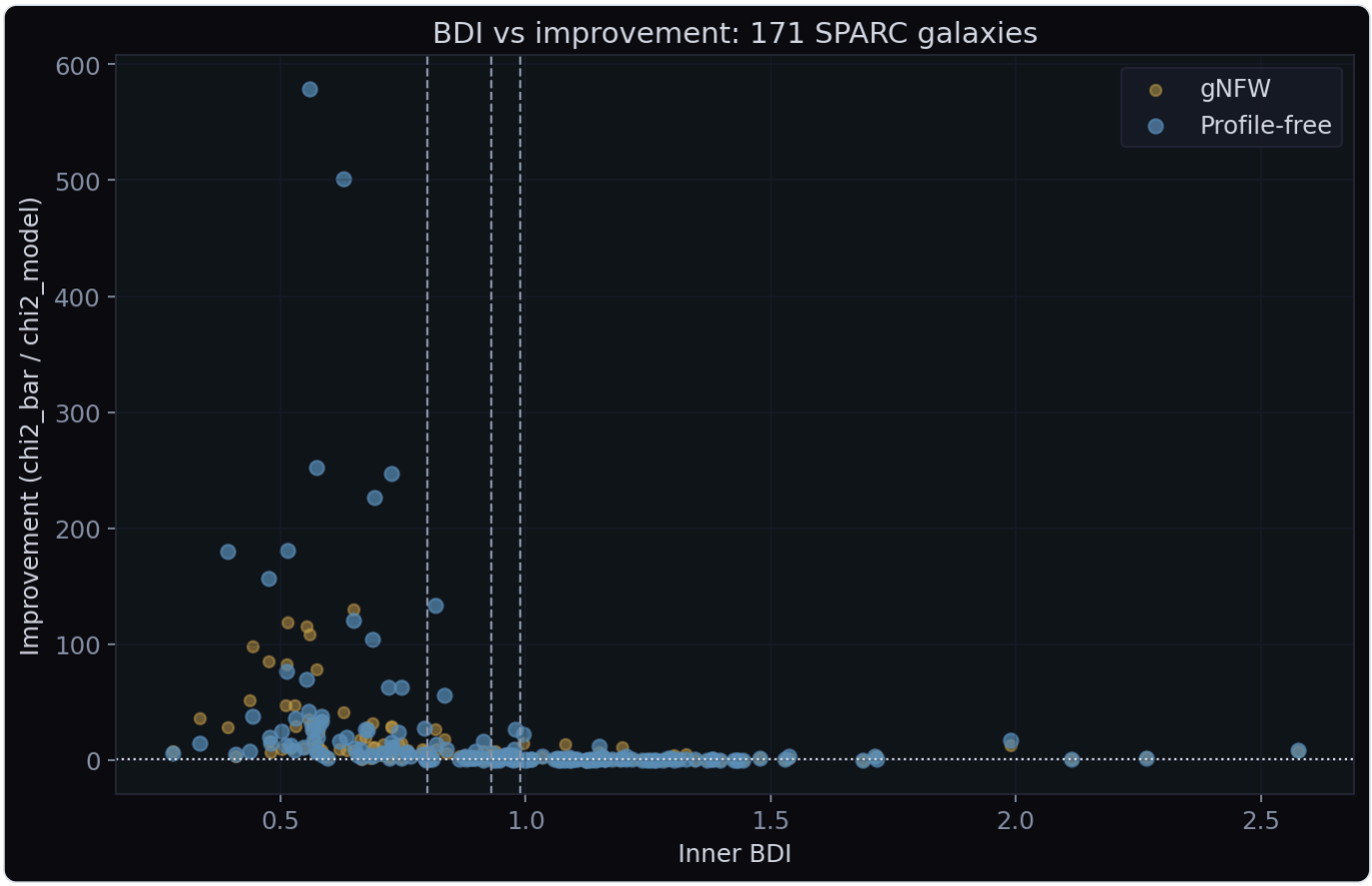
Largest profile-free wins over gNFW

GALAXY	REGIME	INNER_BDI	CHI2_BAR	CHI2_GNFW	CHI2_PF	PF_IMPROVEMENT	DELTA_PF_GNFW	BEST_MOD
UGC02953	>=0.99	1.123	298.95	4.642e+03	983.02	0.304	-3.659e+03	potential_
NGC5055	>=0.99	1.381	593.26	3.481e+03	1.169e+03	0.508	-2.312e+03	potential_
UGC09133	>=0.99	1.093	499.20	2.476e+03	1.148e+03	0.435	-1.328e+03	potential_
NGC2403	>=0.99	0.998	191.66	1.501e+03	630.40	0.304	-870.21	spectral_v
NGC6015	0.93-0.99	0.946	109.96	876.57	241.80	0.455	-634.77	spectral_v

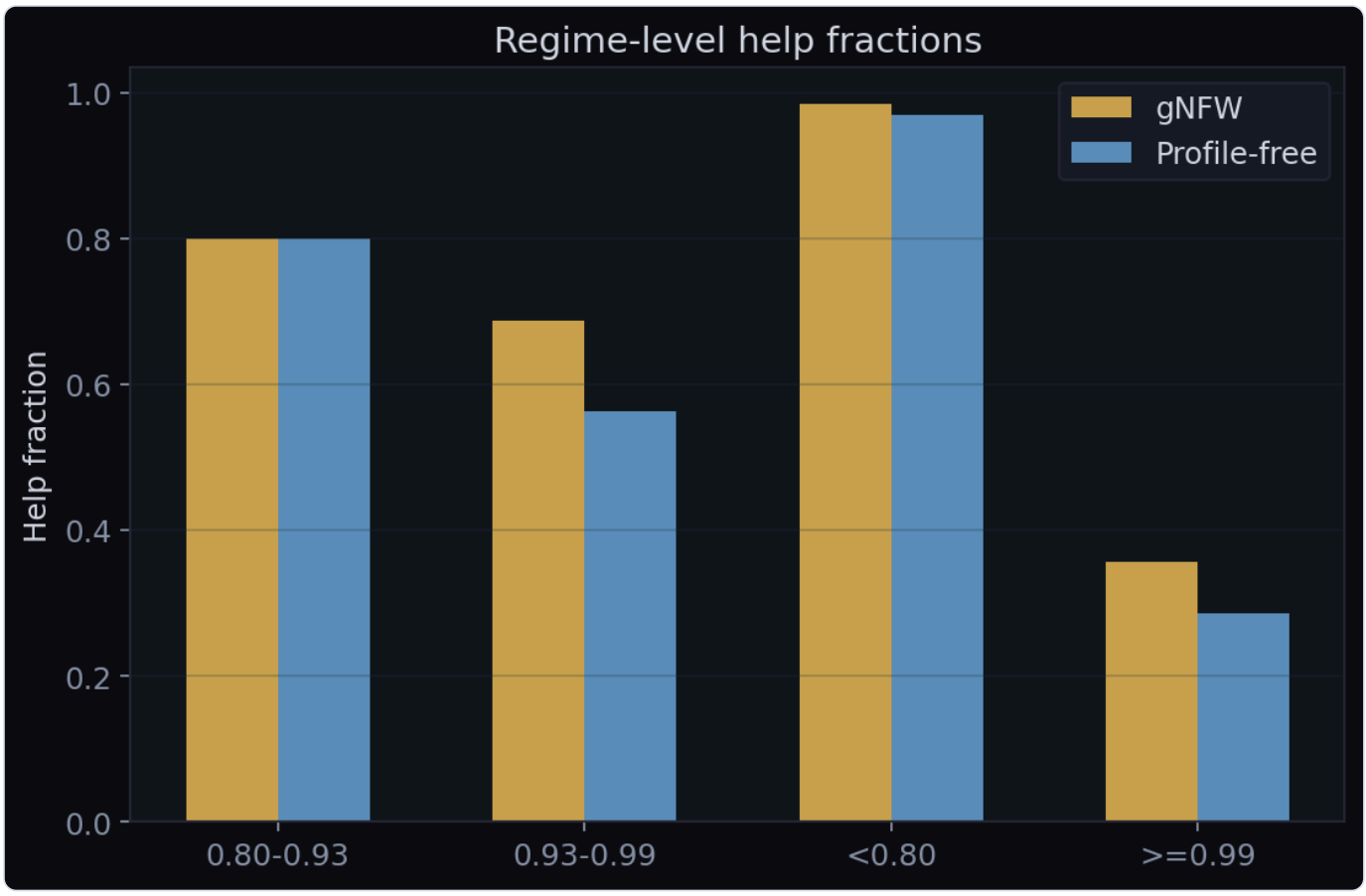
GALAXY	REGIME	INNER_BDI	CHI2_BAR	CHI2_GNFW	CHI2_PF	PF_IMPROVEMENT	DELTA_PF_GNFW	BEST_MOD
NGC2841	0.80-0.93	0.891	427.03	839.16	243.16	1.756	-596.00	potential_
NGC3198	>=0.99	1.106	308.33	799.23	453.77	0.679	-345.45	potential_
NGC6503	>=0.99	1.011	388.96	1.129e+03	786.71	0.494	-342.51	spectral_v
NGC6674	0.80-0.93	0.889	1.187e+03	726.77	488.49	2.431	-238.28	potential_

The hardest cases include systems such as NGC5055 and other high-BDI galaxies already known to be difficult or resolution-stressed in the earlier 50-galaxy analyses. That continuity is scientifically useful: it suggests that the new failures are not random, but connected to the same regime structure already seen at smaller sample size.

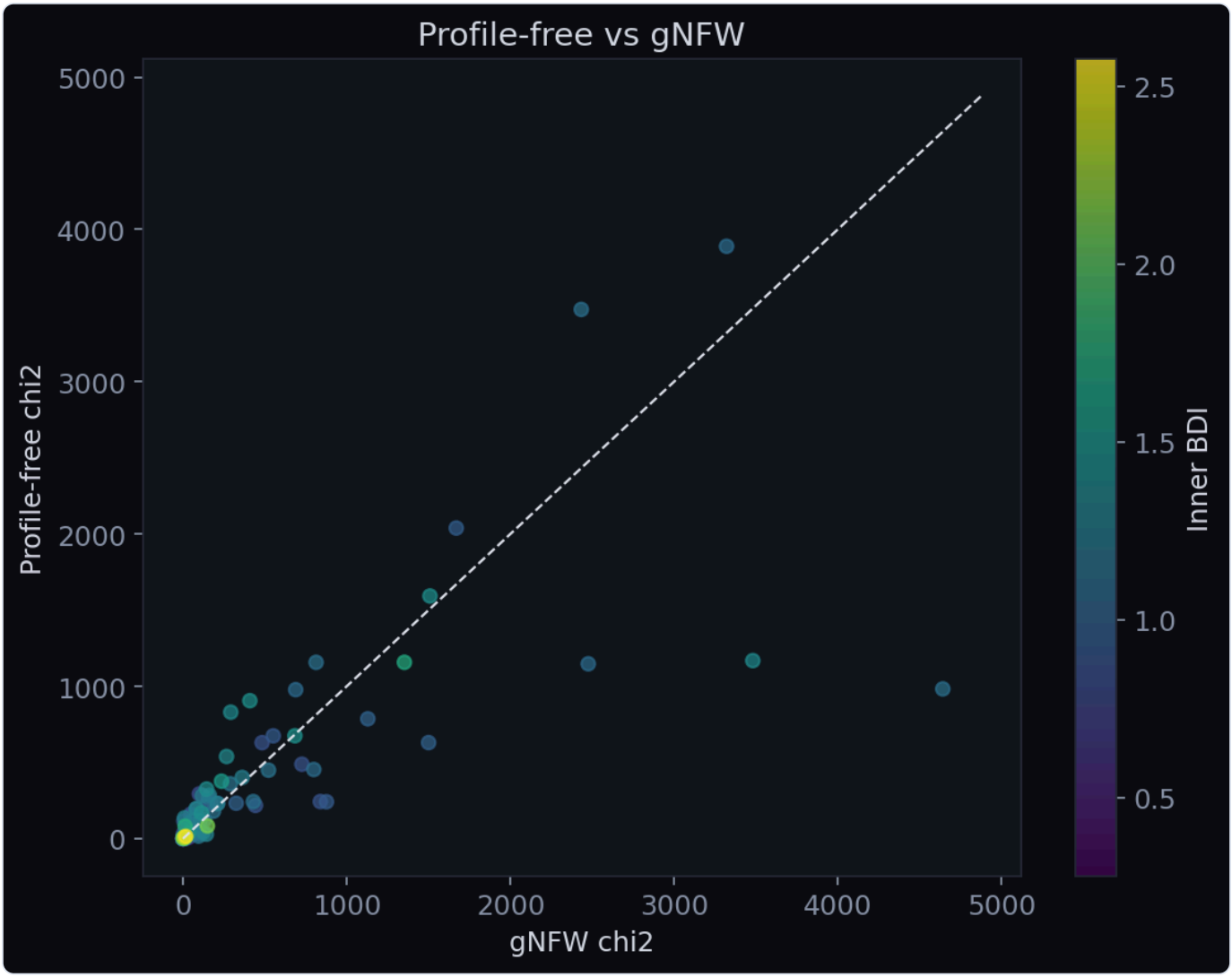
8. All available figures and how to read them



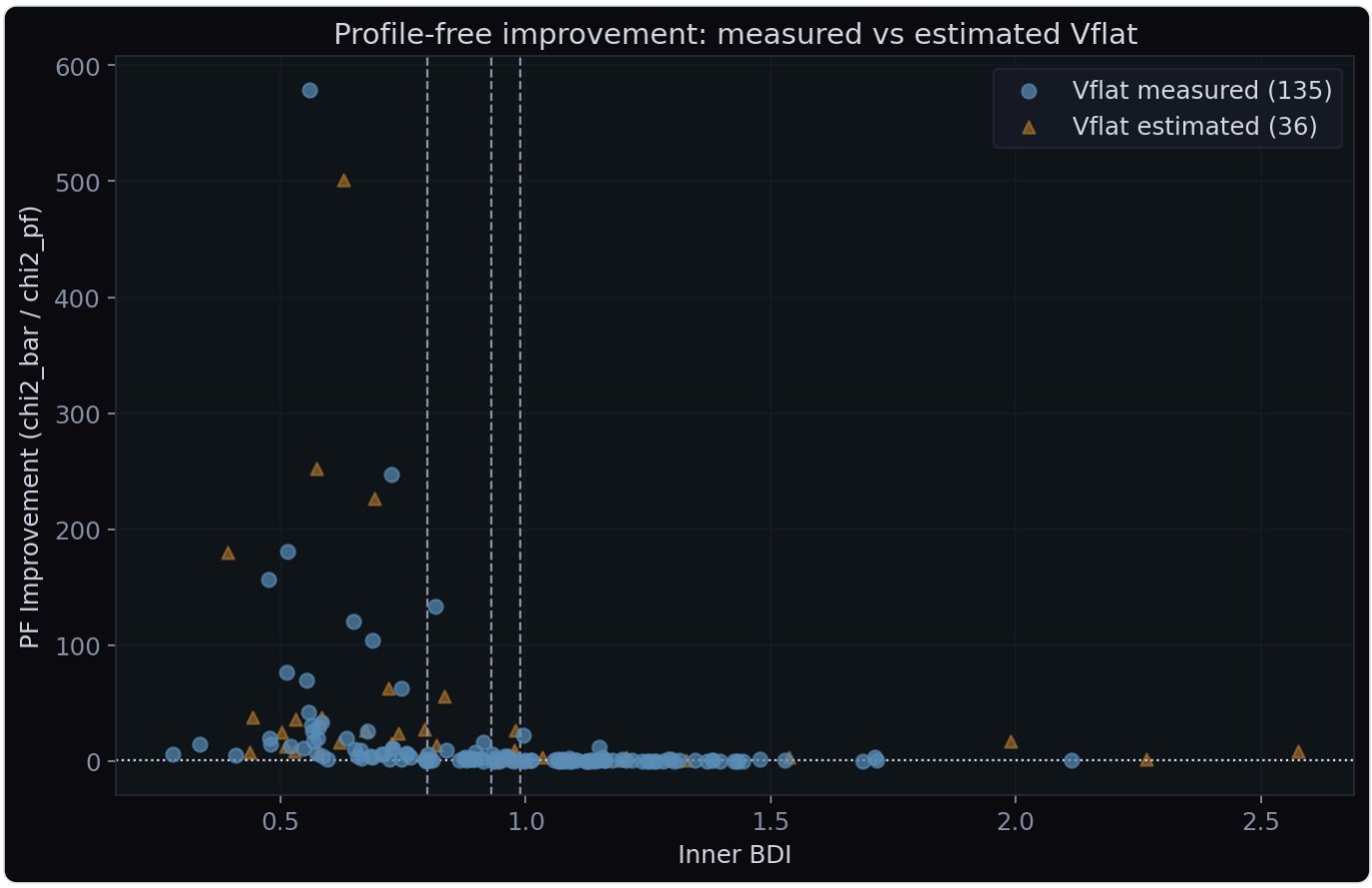
Phase 11D v3: BDI versus profile-free improvement. Core large-sample result. The anti-correlation shows that profile-free improvement is concentrated toward lower inner-BDI systems, consistent with the previously identified applicability-domain structure.



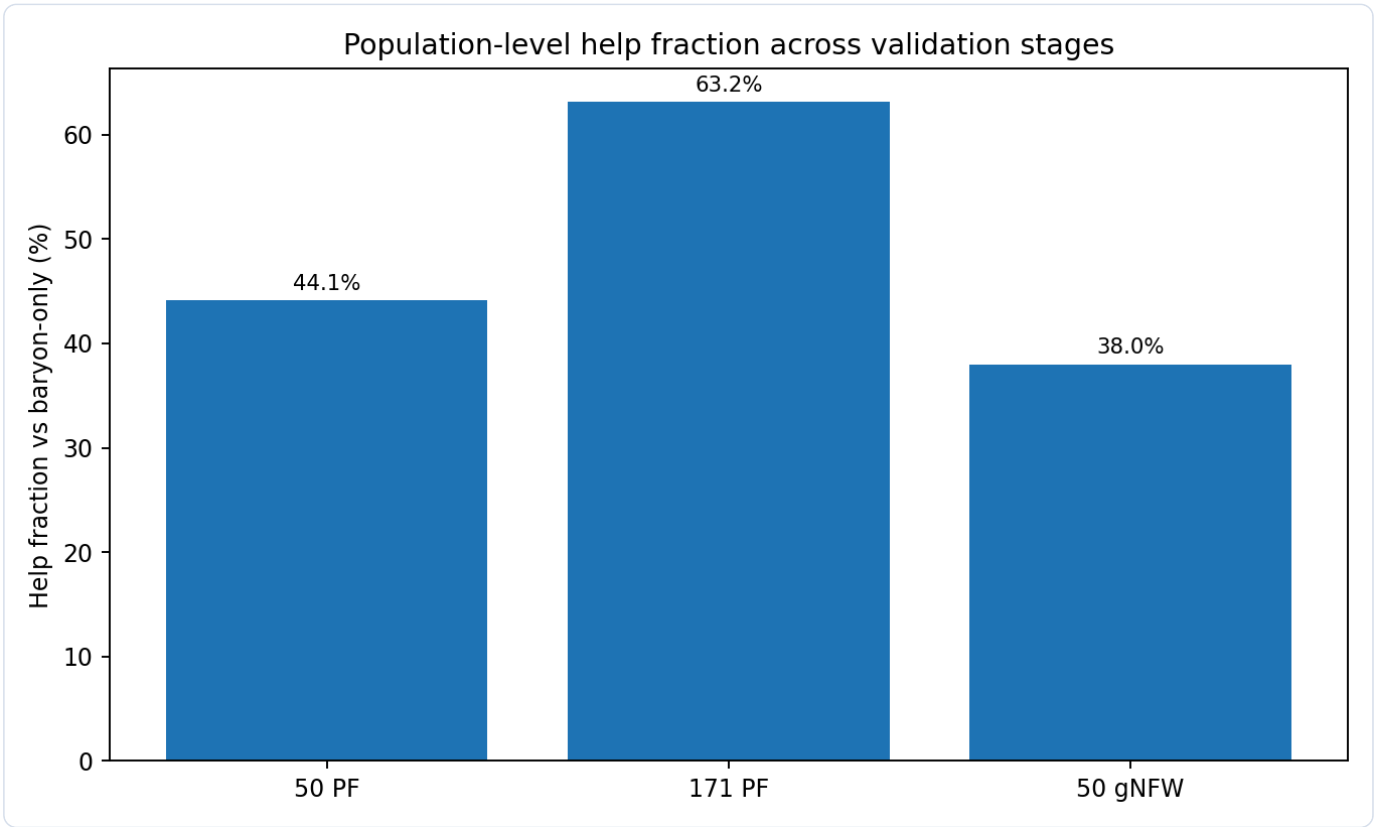
Phase 11D v3: regime help fractions. Directly shows where the profile-free and gNFW branches help relative to baryon-only. The low-BDI regime remains the strongest domain of help, while the ≥0.99 regime remains substantially harder.



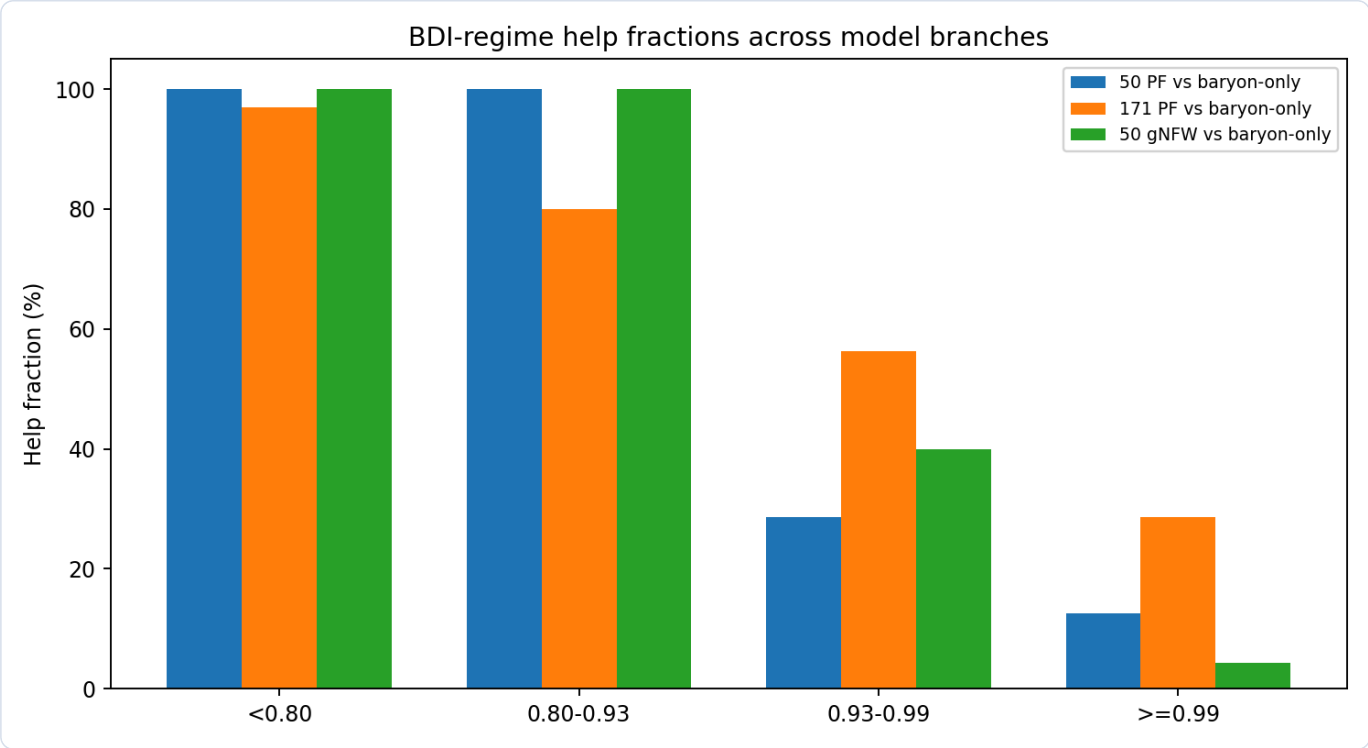
Phase 11D v3: χ^2 profile-free versus χ^2 gNFW. Pairwise comparison against the standard parametric gNFW reference on the same 171-galaxy sample.



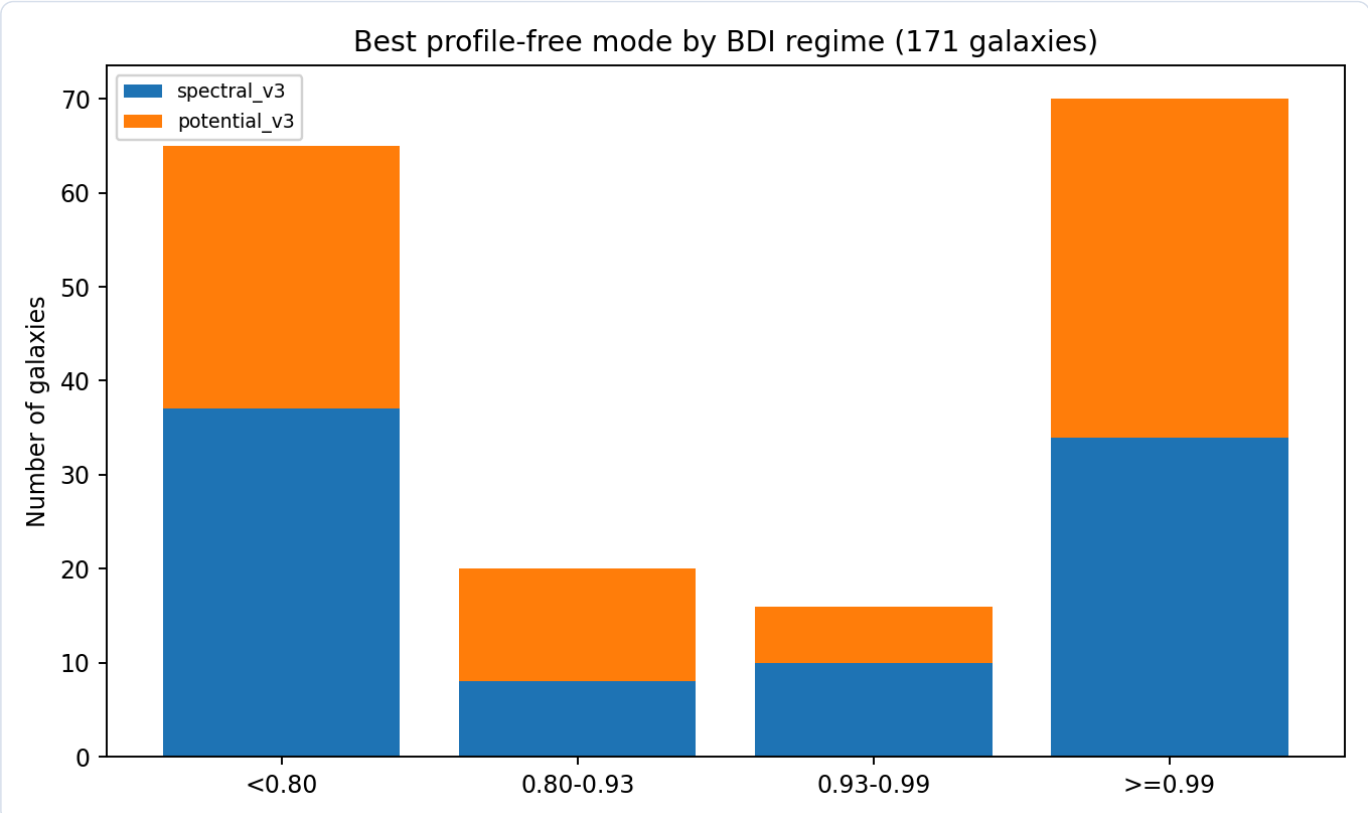
Phase 11D v3: measured versus estimated BDI. Quality-control figure for the mixed measured / estimated-vflat catalogue. Useful for assessing whether the main trend may be driven by the estimated subset.



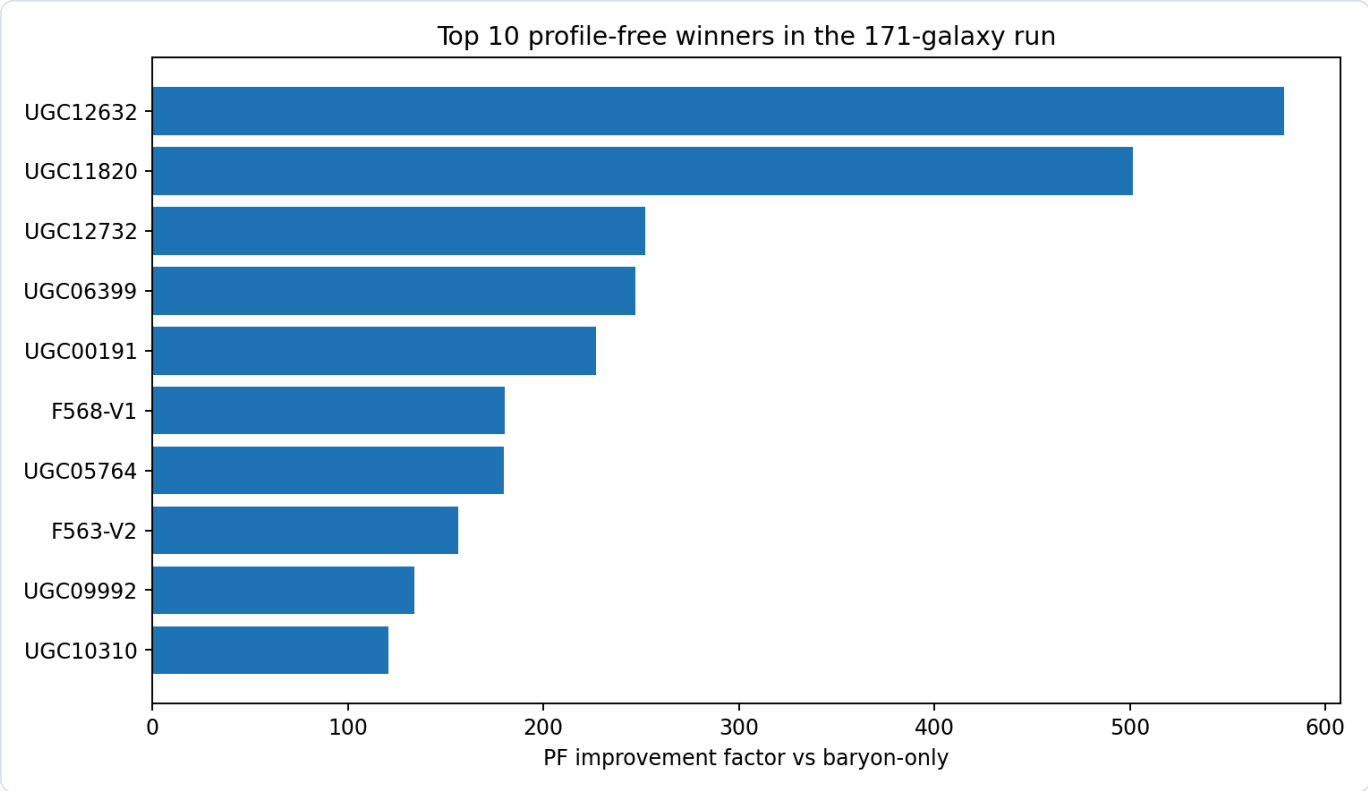
New synthesis figure: help fraction across stages. Compares the fraction of galaxies improved over baryon-only across the Phase 10e parametric branch, the 50-galaxy profile-free pilot, and the 171-galaxy profile-free run.



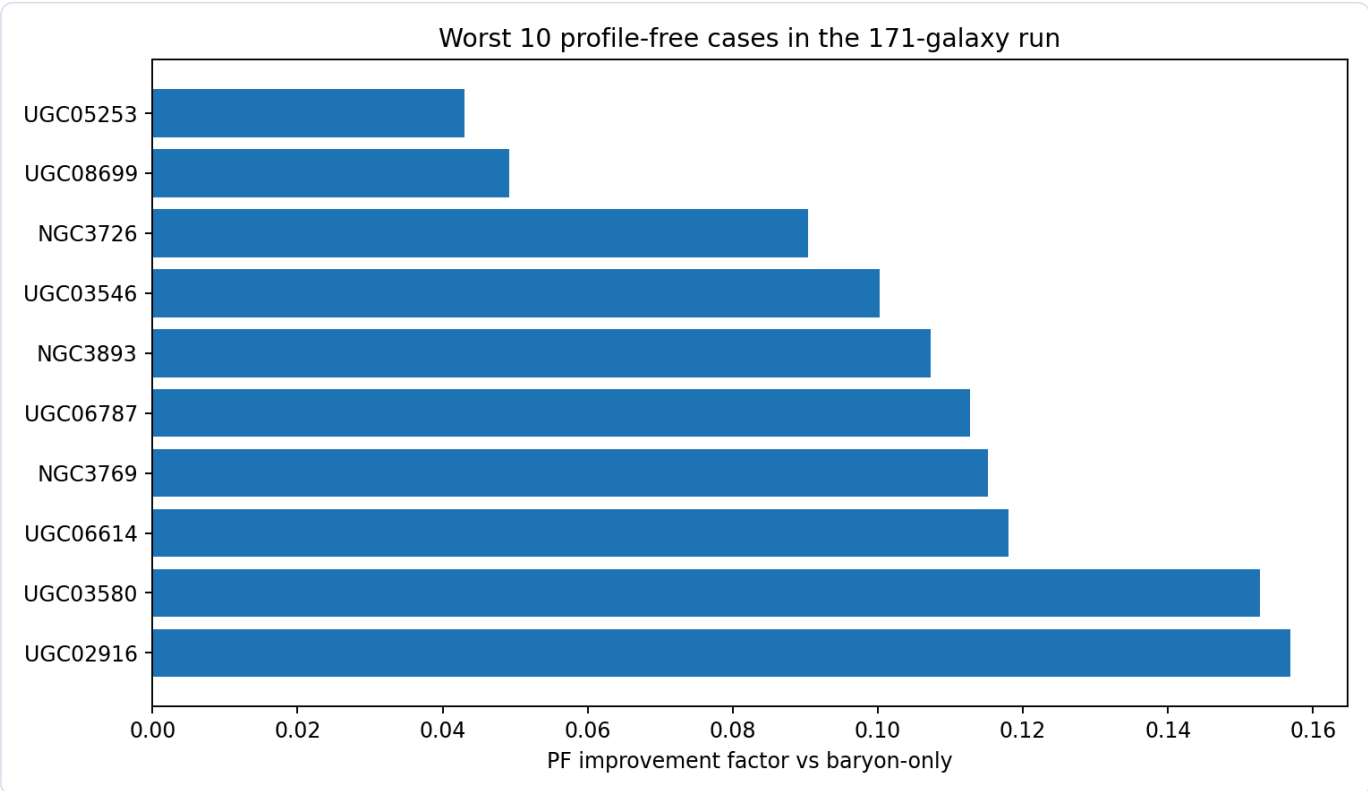
New synthesis figure: regime comparison across branches. Places the 171-galaxy profile-free regime pattern next to the earlier 50-galaxy profile-free pilot and the 50-galaxy gNFW branch.



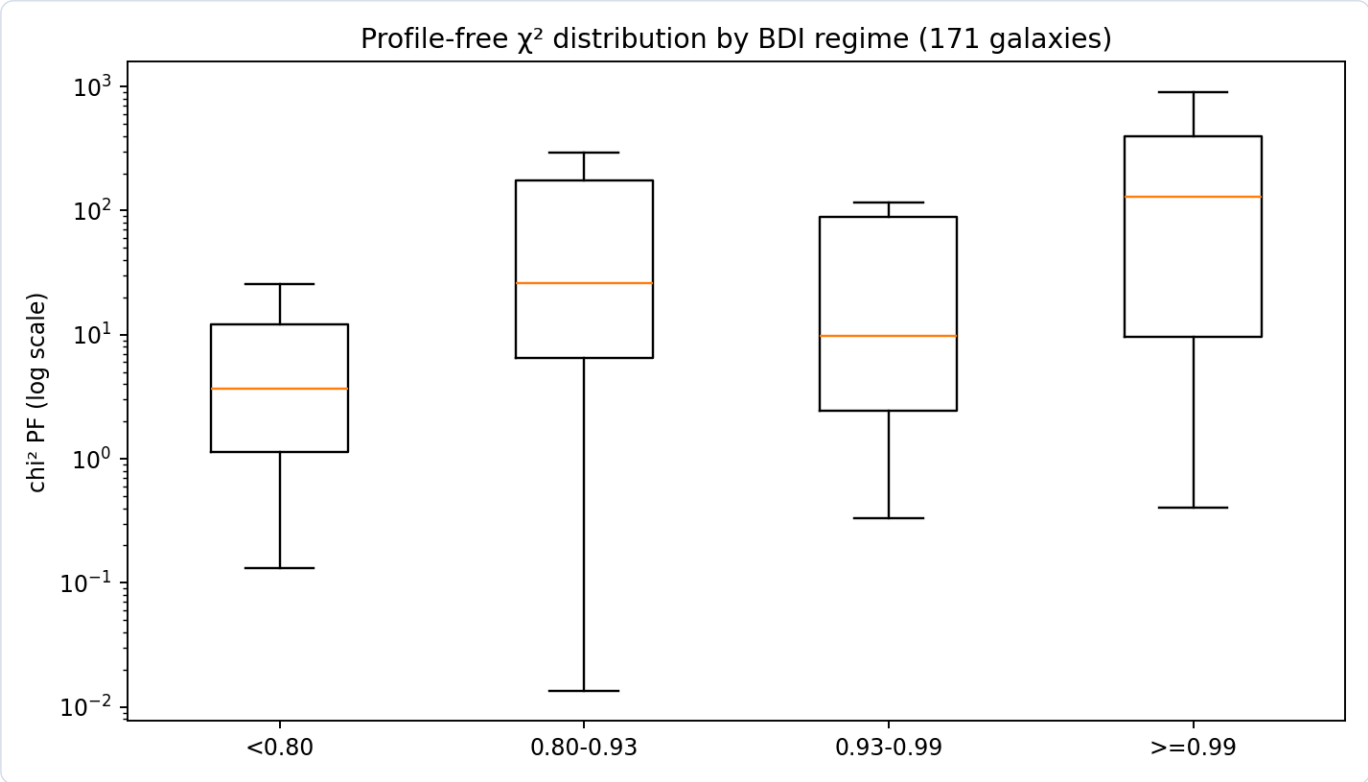
New synthesis figure: best mode by BDI regime. Shows that neither spectral nor potential mapping completely dominates globally; the winning mode distribution varies by regime.



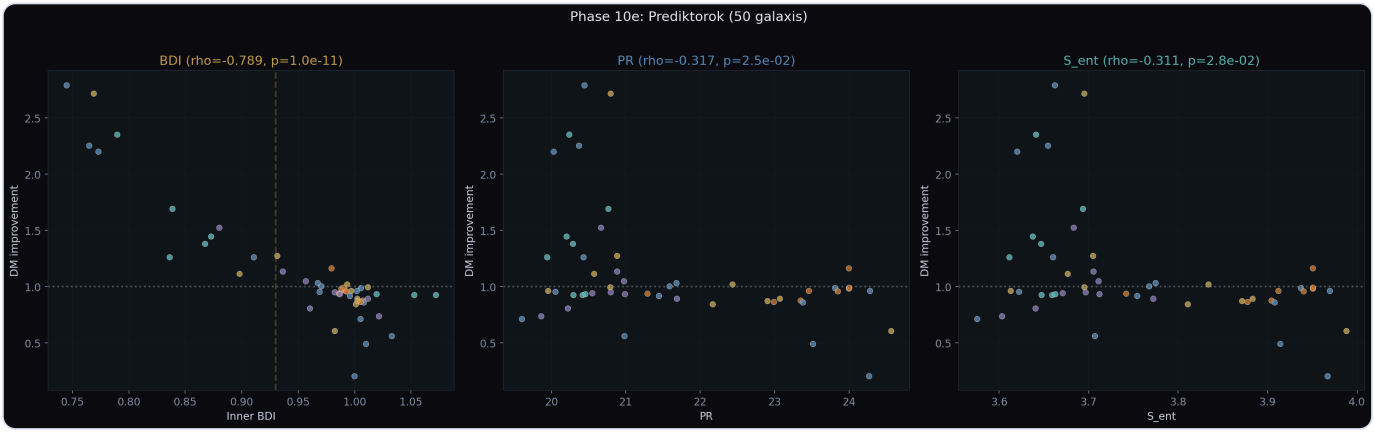
New synthesis figure: top 10 profile-free winners. Largest improvement factors versus baryon-only. Most belong to the low-BDI region and many are gas-rich, low-surface-brightness systems.



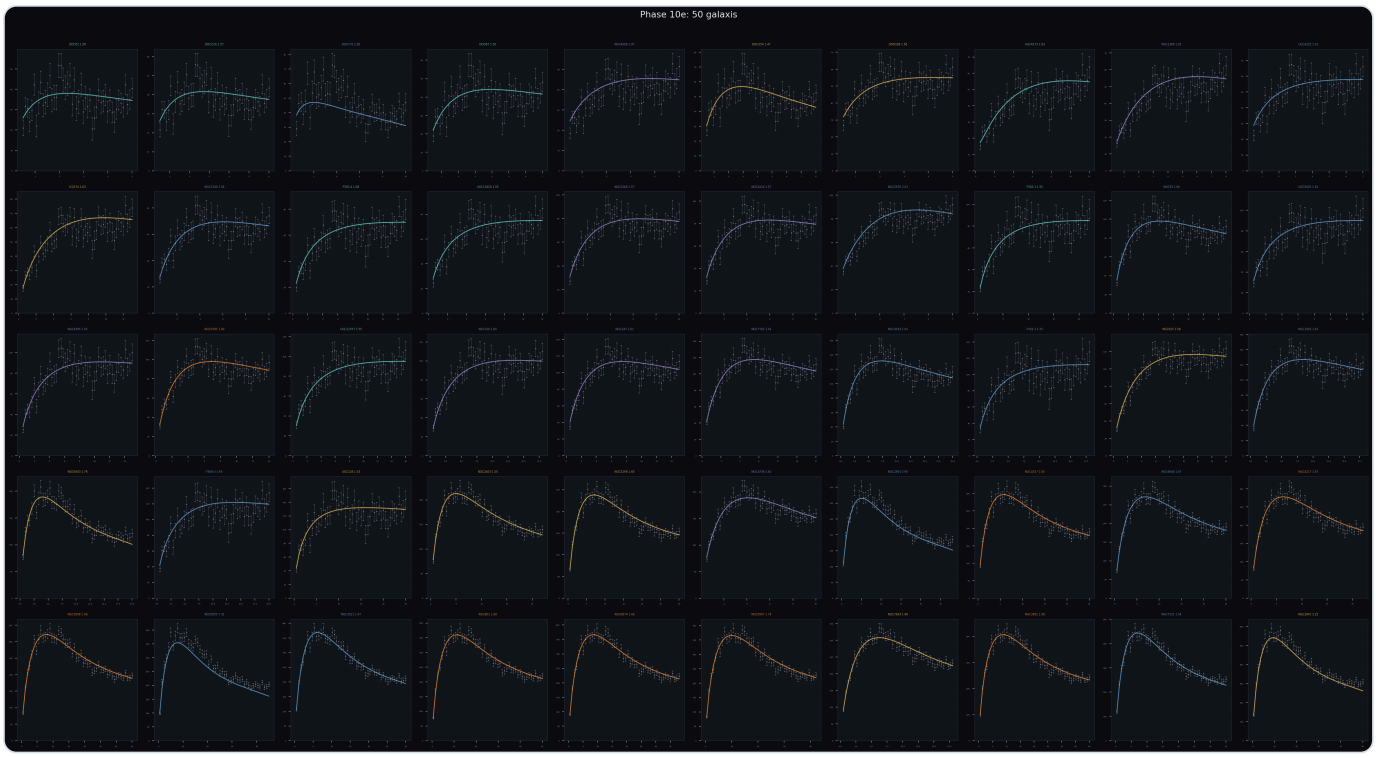
New synthesis figure: worst 10 profile-free cases. Worst systems against baryon-only. These cases are concentrated in high-BDI systems and likely mix true mechanistic limits with baseline/modeling difficulty.



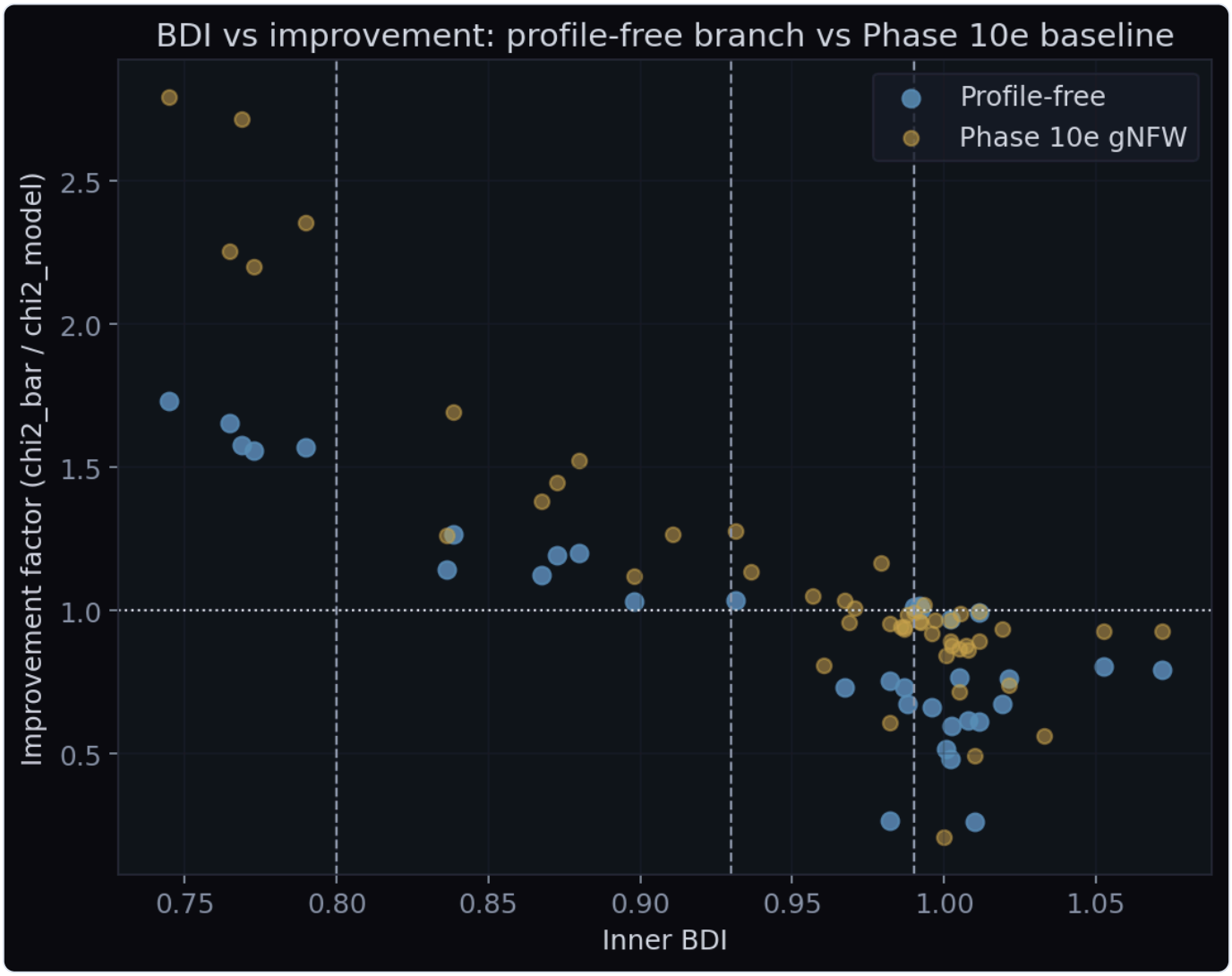
New synthesis figure: χ^2 distribution by regime. Highlights the extreme heterogeneity of the full 171-galaxy run. Low-BDI systems are often dramatically better than baryon-only, but the ≥0.99 regime still contains very hard outliers.



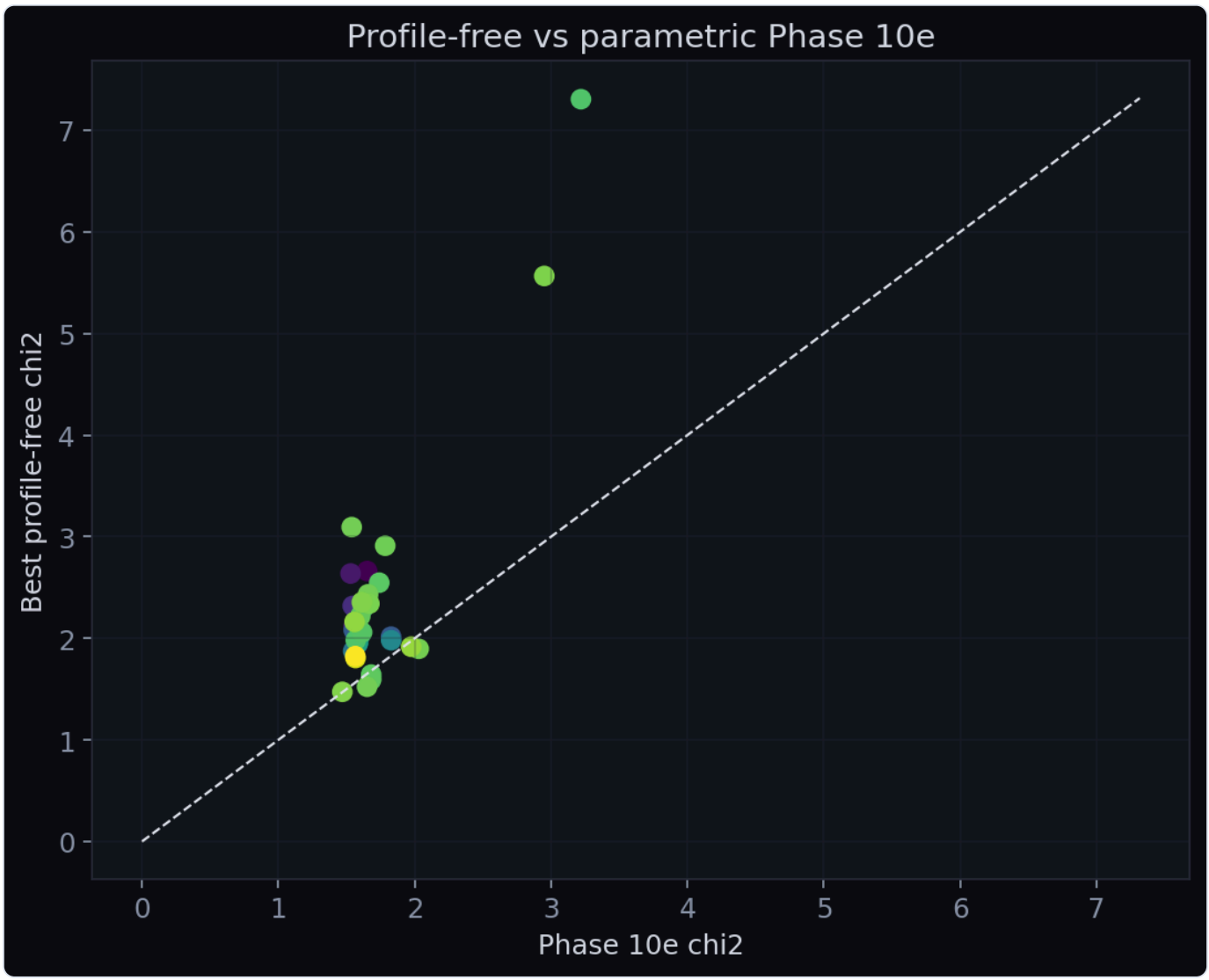
Phase 10e predictors (50-galaxy parametric branch). Earlier 50-galaxy predictor figure showing BDI as primary separator and PR / spectral entropy as weaker secondary predictors.



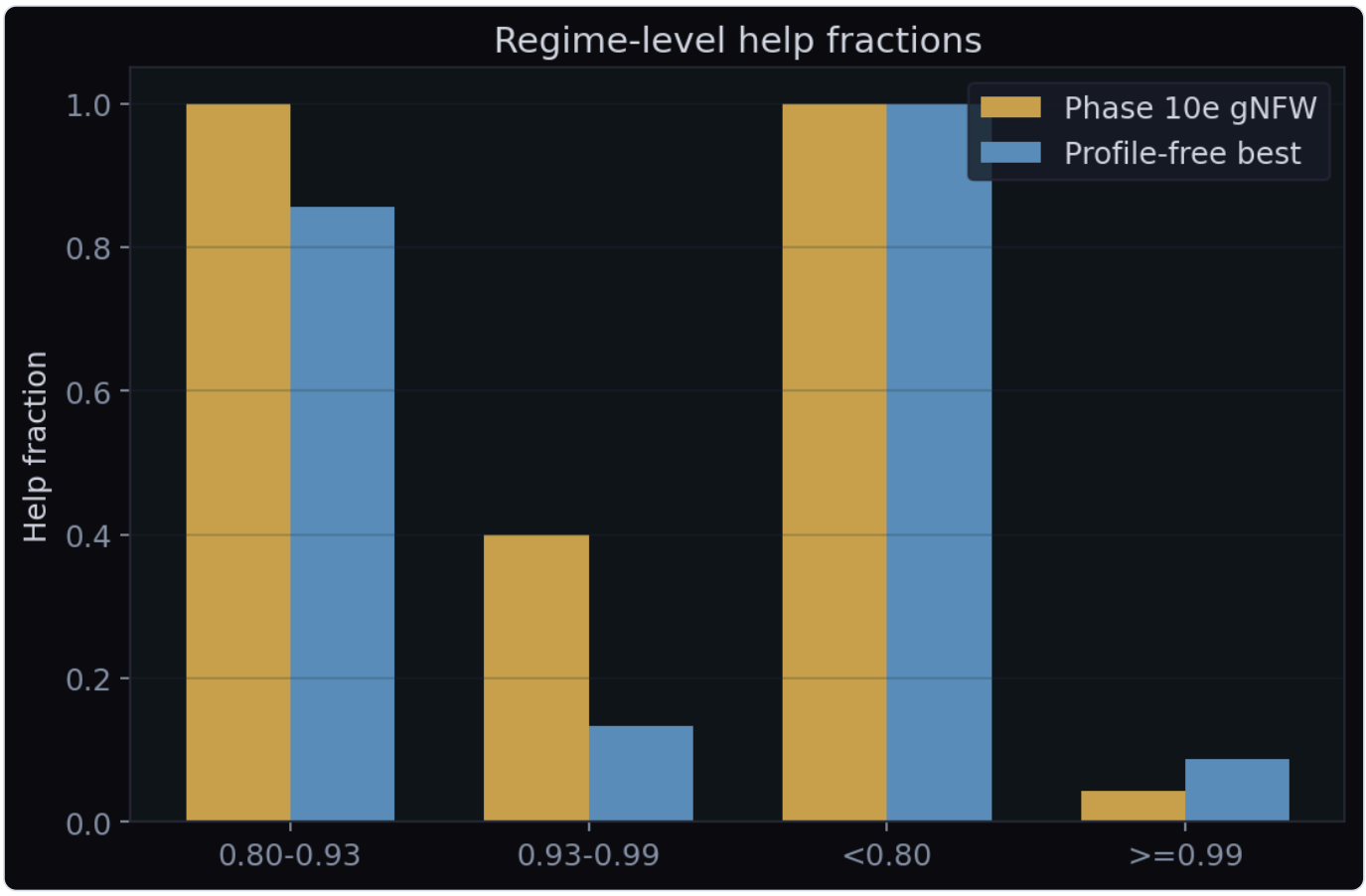
Phase 10e rotation-curve mosaic. Earlier 50-galaxy rotation-curve overview for the parametric ceiling branch.



Phase 11D v2 (50-galaxy profile-free): BDI vs improvement. The first 50-galaxy profile-free pilot already recovered the same qualitative regime structure, although with weaker quantitative performance than the parametric branch.



Phase 11D v2 (50-galaxy profile-free): χ^2 profile-free vs Phase 10e. Direct comparison between the first 50-galaxy profile-free pilot and the 50-galaxy parametric reference.



Phase 11D v2 (50-galaxy profile-free): regime help fractions. Shows that the regime pattern was already present in the 50-galaxy pilot, even before the 171-galaxy run.

---

9. What threshold is actually crossed by the supplied data

Supported by the present evidence. The 171-galaxy profile-free run demonstrates that the BDI-governed applicability-domain structure persists beyond the earlier 50-galaxy validation cohort. The profile-free branch is not merely structurally suggestive: it improves on baryon-only in a majority of the supplied sample and slightly exceeds the gNFW reference in aggregate. In that limited but important sense, the supplied evidence crosses a meaningful barrier between small-sample mechanistic promise and large-sample competitive performance.

Not yet supported. The present evidence does not justify claiming that TAIJITU is already a universally successful galaxy-fit theory, that the current profile-free pipeline has solved the large-galaxy/high-BDI tail, or that the absolute χ^2 scale is already competitive with the best dedicated halo-fitting pipelines on every system.

The strongest careful conclusion is therefore this: the profile-free branch has crossed from a merely mechanistic demonstration into a large-sample competitive branch, but it has not yet eliminated the need for further mapping refinement, baseline harmonisation, and targeted treatment of the hardest high-BDI systems.

---

10. Limitations and uncertainty sources

- **Sample difficulty:** the full 171-galaxy run is substantially harder than the earlier 50-galaxy validation cohort, so χ^2 values are not directly comparable without context.

- **Parsing loss:** 4 of 175 galaxies were lost during parsing. This is not enough to invalidate the result, but it shows the catalogue pipeline is not yet perfectly closed.
- **Mixed V_{flat} quality:** 135 systems use directly measured V_{flat}, while 36 use fallback/estimated values. The measured-vs-estimated BDI figure therefore matters for quality control.
- **Absolute-fit ceiling:** even though profile-free slightly beats gNFW in median χ^2, both branches still have large absolute χ^2 on hard systems.
- **No blind benchmark yet:** the run is large-sample and informative, but it is not yet the final blind probabilistic validation framework envisioned in the broader programme.

---

11. Optimal next steps toward a decisive benchmark

The next steps should follow directly from what the data now show, not from generic model-development instinct. The immediate goal is to turn a threshold result into a decisive benchmark.

1. **Freeze this 171-galaxy result into a reproducible benchmark package.** This run changes the scientific status of the profile-free branch. It therefore deserves a locked benchmark export, a documented manifest, and a stable reference table before further exploratory tuning begins.
2. **Perform full regime-level disaggregation on the 171-galaxy sample.** The headline numbers are already strong, but the scientific interpretation still depends on BDI regime structure. A definitive publication should therefore break out all key metrics by the four BDI regimes, by morphology, by measured-versus-estimated V_{flat}, and by mode winner.
3. **Audit the high-BDI tail separately.** The ≥0.99 regime remains the main failure concentration. This is where the next gain is most likely to come from if profile-free mapping is improved, and it is also where baseline ambiguity and true mechanistic limitation are most likely to be entangled.
4. **Analyse spectral versus potential mode specialisation.** The near-even 89/82 split implies a real scientific question: are these modes tracking different physical sub-populations, or merely compensating for current pipeline imperfections?
5. **Bridge the 50-galaxy and 171-galaxy papers explicitly.** The 50-galaxy parametric branch remains the cleanest ceiling result; the 171-galaxy run is the strongest profile-free result. A publication strategy that explicitly links these two claims will be more convincing than treating them as separate stories.
6. **Only after that, move to hierarchical population modelling and blind probabilistic validation.** The current evidence now justifies Phase 11B v2 and Phase 11A v2 style work, because the profile-free branch is finally informative enough at large sample size to support those more formal statistical steps.

---

12. Appendix: full data tables

▼ Appendix A — 171-galaxy master table used in this manuscript

GALAXY	MORPH	LSB	Q	VFLAT_MEASURED	VF	MB	GAS_FRAC	INNER_BDI	REGIME	CHI2_BAR
UGC00731	Irr	True	1	True	73.30	2.565e+09	0.937	0.280	<0.80	208.96
UGC06667	Scd-Sd	True	1	True	83.80	1.774e+09	0.606	0.336	<0.80	248.27
UGC05764	Irr	True	2	False	10.00	2.593e+08	0.836	0.392	<0.80	223.35
UGCA444	Irr	True	2	True	37.00	9.511e+07	0.937	0.408	<0.80	12.53
UGC05918	Irr	True	2	False	10.00	5.115e+08	0.772	0.437	<0.80	29.16
F565-V2	Irr	True	2	False	10.00	1.209e+09	0.769	0.442	<0.80	51.98
F563-V2	Irr	True	1	True	116.60	4.378e+09	0.659	0.476	<0.80	20.49
F563-1	Sdm-Sm	True	1	True	109.90	5.208e+09	0.817	0.477	<0.80	30.23
UGCA442	Sdm-Sm	True	1	True	56.40	4.198e+08	0.833	0.480	<0.80	174.31
ESO444	Irr	True	2	False	10.00	2.150e+08	0.835	0.503	<0.80	184.35

▶ Appendix B — full profile-free fit table (all rows from the run archive)

▶ Appendix C — mechanistic signature table

▶ Appendix D — parsed SPARC-like catalogue table

▶ Appendix E — baseline table for all parsed galaxies

▶ Appendix F — 50-galaxy profile-free pilot master table

Please consider the environment before printing